Projecting Absolute Invasive Breast Cancer Risk in White Women With a Model That Includes Mammographic Density

Jinbo Chen, David Pee, Rajeev Ayyagari, Barry Graubard, Catherine Schairer, Celia Byrne, Jacques Benichou, Mitchell H. Gail

Background: **To improve the discriminatory power of the Gail model for predicting absolute risk of invasive breast cancer, we previously developed a relative risk model that incorporated mammographic density (DENSITY) from data on white women in the Breast Cancer Detection Demonstration Project (BCDDP). That model also included the variables age at birth of first live child (AGEFLB), number of affected mother or sisters (NUMREL), number of previous benign breast biopsy examinations (NBIOPS), and weight (WEIGHT). In this study, we developed the corresponding model for absolute risk.** ***Methods:*** **We combined the relative risk model with data on the distribution of the variables AGEFLB, NUMREL, NBIOPS, and WEIGHT from the 2000 National Health Interview Survey, with data on the conditional distribution of DENSITY given other risk factors in BCDDP, with breast cancer incidence rates from the Surveillance, Epidemiology, and End Results program of the National Cancer Institute, and with national mortality rates. Confidence intervals (CIs) accounted for variability of estimates of relative risks and of risk factor distributions. We compared the absolute 5-year risk projections from the new model with those from the Gail model on 1744 white women.** ***Results:*** **Attributable risks of breast cancer associated with DENSITY, AGEFLB, NUMREL, NBIOPS, and WEIGHT were 0.779 (95% CI = 0.733 to 0.819) and 0.747 (95% CI = 0.702 to 0.788) for women younger than 50 years and 50 years or older, respectively. The model predicted higher risks than the Gail model for women with a high percentage of dense breast area. However, the average risk projections from the new model in various age groups were similar to those from the Gail model, suggesting that the new model is well calibrated.** ***Conclusions:*** **This new model for absolute invasive breast cancer risk in white women promises modest improvements in discriminatory power compared with the Gail model but needs to be validated with independent data. [J Natl Cancer Inst 2006; 98:1215–26]**

Gail et al. *(1)* used data from the Breast Cancer Detection Demonstration Project (BCDDP) to develop a model for the absolute risk of breast cancer for women in a given age interval. This model, known as the Gail model, included age at menarche (AGEMEN), age at birth of first live child (AGEFLB), number of previous benign breast biopsy examinations (NBIOPS), and number of first-degree relatives (mother or sisters) with breast cancer (NUMREL). We call these standard risk factors. This model was recalibrated to data from National Cancer Institute's (NCI's) Surveillance, Epidemiology, and End Results (SEER) program *(2)*, and the resulting model, called Gail model 2, is available at http://www.cancer.gov/bcrisktool/. This model has been used to design prevention trials, such as the Breast Cancer Prevention Trial *(3)*, and to counsel women about their individual risks *(4)*. Although the model accurately predicts the numbers of cancers observed in subsets of women defined by various combinations of age and risk factors (i.e., it is well calibrated) *(2,5)*, it has been criticized because the distribution of risks in women who develop breast cancer has considerable overlap with the distribution of risks in women who do not develop breast cancer *(5)*. The corresponding low concordance statistic has been described as a lack of discriminatory power *(5)*.

We investigated whether mammographic density could be used to improve the discriminatory power of the model because increased mammographic density has been associated with strongly increased breast cancer risk *(6–8)* and with high attributable risk *(6)*. We defined the variable DENSITY in terms of the average percent dense area of the two breasts on craniocaudal mammograms. We previously obtained measurements of the variable DENSITY in 7251 women from the BCDDP and determined that the reader reliability of these measurements was sufficient to warrant their inclusion in risk models *(9)*.

As a first step toward developing a model of absolute risk that includes DENSITY, we developed a relative risk model (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). The technical challenge was to estimate efficiently the effects of standard risk factors, which were measured on large numbers of women, and the effects of DENSITY, which was measured only on subsets of the women for which all of the other risk factors had been measured (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). Chen et al. (unpublished results) examined several possible relative risk models. A parsimonious model that captured the most important effects included AGEFLB, NBIOPS, NUMREL, DENSITY, and weight (WEIGHT). Unlike the original model of Gail et al. *(1)*, this model did not include AGEMEN, an interaction between AGEFLB and NUMREL, or an interaction between NBIOPS and whether age equaled or exceeded 50 years because these terms were not statistically significantly associated with breast cancer risk in the multivariable model. Chen et al.

Affiliations of authors: Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA (JC); Information Management Services, Rockville, MD (DP, RA); Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD (BG, CS, MHG); Cancer Genetics and Epidemiology Program, Georgetown University Lombardi Comprehensive Cancer Center, Washington, DC (CB); Medical School, Rouen University Hospital, Inserm U 657, Rouen, France (JB).

Correspondence to: Mitchell H. Gail, MD, PhD, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892 (e-mail: gailm@mail.nih.gov).

See "Notes" following "References."

DOI: 10.1093/jnci/djj332

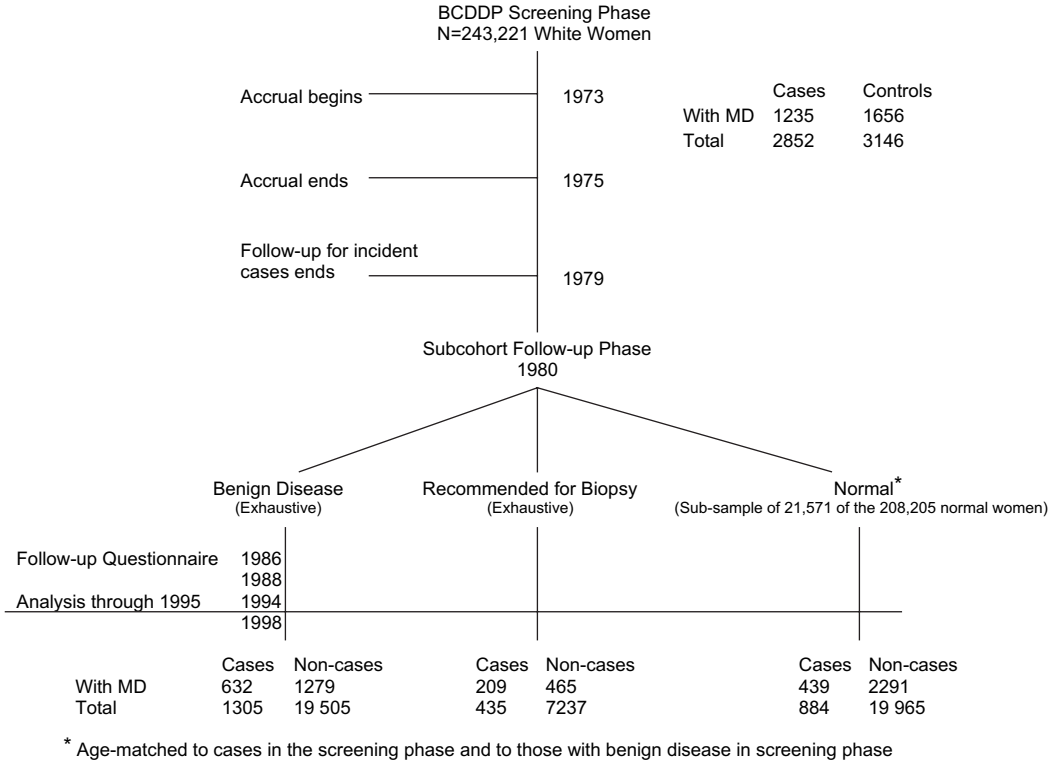

Fig. 1. Participation and mammographic density measurements in the Breast Cancer Detection Demonstration Project (BCDDP). MD refers to mammographic density. Cases refer to women with breast cancer. Noncases refer to women without breast cancer, some of whom were selected as control subjects.

(unpublished results) analyzed the concordance statistic from the new model and found increases in concordance ranging from 0.01 to 0.09 in the seven 5-year age groups studied. It seemed worthwhile, therefore, to develop the corresponding model for absolute risk.

In this study, we developed a model to estimate absolute invasive breast cancer risk that was based on age and on DENSITY, WEIGHT, AGEFLB, NBIOPS, and NUMREL. We also compared estimates from this model of absolute risk, which included DENSITY, with those from Gail model 2 *(2)*. Calculations of the variance needed to put confidence intervals (CIs) on the absolute risk projections are described in the Appendix.

Methods and Data Sources

Strategy to Estimate Absolute Risk

The components *(1)* needed to project absolute risk are the relative risk model, the baseline hazard rate for invasive breast cancer at age t [$h_1(t)$], and the corresponding hazard of mortality from non–breast cancer causes [$h_2(t)$]. We used estimates of $h_2(t)$ from the National Center for Health Statistics for January 1, 1996, through December 31, 2000 *(10)*. We obtained $h_1(t)$ by multiplying SEER hazard rates *(11)* from January 1, 1996, through December 31, 2000, for invasive breast cancer, $h_1^*(t)$, by [1 − AR(t)], where AR(t) is the attributable risk for the entire relative risk model at age t. It is a technical challenge to estimate the attributable risk because the joint distribution of all the risk factors in the general population is not available. We, therefore, obtained an estimate of the joint distribution of age, WEIGHT, AGEFLB, NBIOPS, and NUMREL from the National Health Interview Survey (NHIS) for 2000 *(12)*. We completed our model for the joint distribution of these factors with DENSITY by using BCDDP data to model the conditional distribution of DENSITY given age, WEIGHT, AGEFLB, NBIOPS, and NUMREL. From resulting estimates of the relative risk model (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results), $h_1(t)$, and $h_2(t)$, we computed absolute risk from formulas 5 and 6 in Gail et al. *(1)*.

Estimating the Relative Risk Function That Incorporates Mammographic Density

The study design of the BCDDP, sampling to obtain information on factors in the Gail model, and sampling of data on mammographic density have been reported previously *(9)* (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). Chen et al. (unpublished results) described how to account for the patterns of missing data when estimating the relative risk function of these variables and how to calculate the variance of the relative risk function. In brief, the BCDDP recruited 284 780 women, including 243 221 white women, from January 1, 1973, through December 31, 1975, and followed the women with annual screening for up to 5 years (Fig. 1). More than 99% of the women were between the ages of 35 and 74 years at recruitment, with a median age of 50 years. The period of BCDDP from January 1, 1973, through December 31, 1979, was referred to as the screening phase of the study. At the end of the screening phase, follow-up was extended for all women who had developed breast cancer; all those who had a benign breast biopsy during the screening phase; all those who were recommended for biopsy examination during the screening phase but did not have it; and a subset of those remaining "normal" women who had not had a breast cancer diagnosis, a breast biopsy, or a recommendation for breast biopsy during the screening phase (Fig. 1). Normal subjects were selected by frequency matching on age, date of entry into the BCDDP screening phase, race, center, and length of participation in screening to women

who developed breast cancer or who had a biopsy in the screening phase. Follow-up continued through December 31, 1998; we used follow-up data through December 31, 1995. We refer to the period from January 1, 1980, through December 31, 1995, as the follow-up phase of BCDDP and to the three groups of women (i.e., women who had biopsy examinations, women recommended for a biopsy examination but who did not receive it, and normal women) as subcohorts.

To study the risk factors for breast cancer, a nested case–control study was conducted within the screening phase in 1979, and data were collected on standard risk factors, including age at menarche, age at birth of first live child, number of breast biopsy examinations, and number of first-degree relatives (mother or sisters) with breast cancer (Fig. 1). Women who had either invasive or in situ breast cancer were defined as case patients. Control subjects for this study were selected from those who did not have a breast cancer diagnosis, a breast biopsy, or a recommendation for breast biopsy during the screening phase and were matched to case patients on age at entry in 5-year intervals, race, study center, 6-month calendar time of screening, and length of follow-up (at least as long as that of the case patient). Gail et al. *(1)* used data on the four standard risk factors and age from the white women in this study to estimate relative odds parameters with a logistic regression model. These standard risk factors were available for most women in the follow-up phase. Of the 21 243 women in the benign biopsy subcohort, 7773 women in the recommended-for-biopsy subcohort, and 21 629 women in the normal subcohort, we excluded 433, 101, and 780 women, respectively, who did not have data on WEIGHT and AGEFLB. There were no women with missing NBIOPS or NUMREL data, which were updated during the follow-up phase. We were able to obtain DENSITY measurements from mammograms at baseline in the period from January 1, 1973, through December 31, 1975 for approximately half the case patients and control subjects in the screening phase. A nested case–control sample was selected from each subcohort in the follow-up phase, and an effort was made to retrieve screening-phase baseline mammographic images. DENSITY measurements were obtained for approximately half of the case patients and control subjects in these nested case–control studies (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results) (Fig. 1).

Chen et al. (unpublished results) used a logistic regression model to describe the odds ratios (ORs) of breast cancer for the screening-phase data and a piecewise exponential model to estimate the corresponding hazard ratios for data from the follow-up phase. We use the term relative risk to refer to the relative hazard obtained from follow-up phase data and the odds ratios from screening-phase data, which closely approximate the relative hazard (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). Chen et al. developed efficient statistical methods to estimate the relative risk parameters by use of data from all subjects, including those without DENSITY measurements, and they analyzed the screening-phase data and the three follow-up subcohorts separately. Summary log relative risks were obtained from the four sets of log relative risk estimates by multivariate-weighted least squares, in which each component log relative risk vector was multiplied by the inverse of the corresponding covariance matrix and then multiplied by the inverse of the sum of the four covariance matrix inverses.

The final log relative risk model included main effects in five continuous variables: age at birth of first live child (AGEFLB), coded as 0, 1, 2, or 3 for ages younger than 20 years, 20–24 years, 25–29 years or nulliparous, or older than 29 years, respectively; number of biopsies (NBIOPS), coded as 0, 1, or 2 for zero, one, or more than one biopsy examination, respectively, at the time the case patients and control subjects were interviewed; number of affected relatives (NUMREL), coded as 0, 1, or 2 for zero, one, or more than one sister or mother with breast cancer at the time the case patients and control subjects were interviewed; body weight (WEIGHT), coded as 0, 1, 2, 3, 4, or 5 for the weight ranges 100 or less, 101–125, 126–150, 151–175, 176–200, and more than 200 pounds at entry in BCDDP for case patients and control subjects in the screening phase and at the beginning of follow-up otherwise; and DENSITY, coded as 0, 1, 2, 3, or 4 for mammographic density measurements in the ranges 0%, 1%–24%, 25%–49%, 50%–74%, and 75%–100% on the initial screening-phase mammograms. The baseline density 0% corresponds to women with no dense breast tissue. We measured mammographic density in percent by measuring the percentage of breast area outlined on a cranial–caudal image that was dense for each breast. The mammographic density in percent was defined as the average of the two percentages of dense area for the two breasts. The natural logarithm of the relative risk is given as described by Chen et al. (unpublished results) by

$$0.158(\text{AGEFLB}) + 0.256(\text{NBIOPS}) + 0.444(\text{NUMREL}) + 0.216(\text{WEIGHT}) + 0.333(\text{DENSITY}). \quad [1]$$

Thus, for a woman with an AGEFLB of 2, NBIOPS of 1, NUMREL of 1, WEIGHT of 3, and DENSITY of 2, the natural logarithm of the relative risk is $10.3 = \exp[(0.158 \times 2) + (0.256 \times 1) + (0.444 \times 1) + (0.216 \times 3) + (0.333 \times 2)]$ times the hazard of a woman of the same age with all risk factors at the baseline level of zero. We denote coefficients in Eq. 1 by the vector β. Unreported analyses showed that the effect of WEIGHT was very similar when model 1 was fitted separately to women younger than 50 years old and to women 50 years old or older. We present factors (Table 1), corresponding to terms in Eq. 1, that can be used to calculate relative risks easily.

Estimating the Baseline Age-Specific Breast Cancer Hazard Rate [$h_1(t)$] and the Hazard of Death From Causes Other Than Breast Cancer [$h_2(t)$]

We used mortality data from the National Center of Health Statistics *(10)* from January 1, 1996, through December 31, 2000, to estimate $h_2(t)$, and we assumed that $h_2(t)$ was known without error and did not vary with the risk factors in our model. To estimate $h_1(t)$ as in Gail et al. *(1)*, we used the formula

$$h_1(t) = [1 - \text{AR}(t)]h_1^*(t), \quad [2]$$

where $h_1^*(t)$ is the composite invasive breast cancer incidence rate available from the SEER program *(11)* between January 1, 1996, through December 31, 2000; AR(t) is the population attributable risk for women of age t; and $h_1^*(t)$ is assumed to be known without error. The challenge was to estimate AR(t).

Because AR(t) changes very little with age, we assumed that AR(t) was a constant for women younger than 50 years old and possibly another constant for women 50 years old or older. Let

Table 1. Factors for computing combined relative risk*

Risk factor category	Risk factor code	Relative risk factor
A. Age at birth of first live child, y	AGEFLB	
<20	0	1.00
20–24	1	1.17
25–29 or nulliparous	2	1.37
≥30	3	1.61
B. No. of biopsy examinations	NBIOPS	
0	0	1.00
1	1	1.29
≥2	2	1.67
C. No. of first-degree female relatives with breast cancer	NUMREL	
0	0	1.00
1	1	1.56
≥2	2	2.43
D. Weight, pounds	WEIGHT	
<100	0	1.00
100–125	1	1.24
126–150	2	1.54
151–175	3	1.91
176–200	4	2.37
>200	5	2.94
E. Mammographic density, %	DENSITY	
0	0	1.00
1–24	1	1.40
25–49	2	1.95
50–74	3	2.72
75–100	4	3.79
F. Atypical hyperplasia (AH)		
No biopsies		1.00
At least one biopsy and no AH in any biopsy		0.93
No AH found; hyperplasia status unknown for ≥1 biopsy		1.00
AH found in at least one biopsy		1.82

*The combined relative risk is obtained by multiplying the relative risk factors corresponding to each of the categories A, B, C, D, E, and F. If we consider a woman with age at birth of first live child (AGEFLB) = 2, number of previous benign breast biopsy examinations (NBIOPS) = 1, number of affected mother or sisters (NUMREL) = 1, weight (WEIGHT) = 3, and mammographic density (DENSITY) = 2 and assume that it is unknown whether atypical hyperplasia was present from the biopsy examination (factor *F*), then from this table, her combined relative risk is 1.37 × 1.29 × 1.56 × 1.91 × 1.95 × 1.00 = 10.3.

the risk factors be described by vector $X' = (X_1, X'_2)$, where X_1 = (DENSITY) and X'_2 = (AGEFLB, NBIOPS, NUMREL, WEIGHT). Let $p(x|I_t)$ be the population mass function for the risk factors X given age is in interval I_t, where I_t corresponds to $t < 50$ years or $t \geq 50$ years. In the analyses that follows, we use quantities such as $p(x|I_t)$ and its estimate interchangeably. Letting $r(x) = \exp(\beta^T x)$ denote the relative risk at $X = x$ compared with $X = 0$, we computed $AR(I_t)$ from

$$AR(I_t) = 1 - \frac{1}{\sum_x p(x|I_t) r(x)}. \qquad [3]$$

Ideally, we would have survey data from the general population on the distribution of X, but such data on DENSITY are not available. We obtained data on the probability mass distribution of X_2 in the general population from the NHIS. The NHIS is based on a complex weighted cluster sampling design *(12)*. Design-based consistent estimates and variance calculations must take this design into account. We used the program SUDAAN *(13)* for this purpose. We indexed the eight age intervals (<45, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, ≥75 years) by $i = 0, 1, 2, \ldots 7$. From NHIS, we estimated the mass distributions $p(x_2, t = i)$, from which conditional mass functions such as $p(t = i|x_2, I_t)$ and $p(x_2|I_t)$ were computed.

To obtain the required conditional joint probability mass function, $p(x|I_t)$, we assumed that the conditional distribution of X_1 for a given age interval $t = i$ and $X_2 = x_2$ was the same in the BCDDP population as in the general population. From the data from control subjects with DENSITY measurements in each of the three subcohorts, we separately estimated $p(x_1|x_2, t, I_t)$. We estimated $p(x_1|x_2, t, I_t)$, instead of directly estimating $p(x_1|x_2, I_t)$ because age is strongly associated with DENSITY *(9)*; thus, inclusion of age may lead to more precise prediction of DENSITY. More importantly, the normal subcohort was sampled by frequency matching on age in 5-year intervals. Its age distribution is not necessarily equal to that of the general population; therefore, naive estimation without stratifying on age would not be expected to yield consistent estimates of $p(x_1|x_2, I_t)$ for the general population. To accommodate the large number of cells defined by x_2 and t, we further modeled $p(x_1|x_2, t = i, t \geq 50)$ and $p(x_1|x_2, t = i, t < 50)$ by the same polytomous logistic regression model, which included trend effects on t and components of x_2. This assumption of a common logistic model is justified by the good fit of the linear trend model in t. The model we fit was

$$p_\gamma(X_1 = c|x_2, t) = e^{\gamma_{0c} + \gamma_c x_2 + \delta_c t} \left[1 + \sum_{k=1}^{4} e^{\gamma_{0k} + \gamma_k x_2 + \delta_k t} \right]^{-1},$$

$$c = 1, 2, 3, 4,$$

$$p_\gamma(X_1 = 0|x_2, t) = \left[1 + \sum_{k=1}^{4} e^{\gamma_{0k} + \gamma_k x_2 + \delta_k t} \right]^{-1}, \quad c = 0. \qquad [4]$$

In Eq. 4, γ_{0c} is an intercept for DENSITY level c, where $c = 1$, 2, 3, or 4. For most of the women in the follow-up subcohorts, the values of X_2 were not available when DENSITY was measured at the beginning of the screening phase. Instead, we used the values determined at the beginning of the follow-up phase in 1980 to predict X_1 from Eq. 4.

The components of γ_c are the coefficients given separately in Table 2 for the three subcohorts. Such a coefficient estimates the change in the natural logarithm of a relative probability from a unit increase in the corresponding covariate in X_2. The relative probability is the ratio of the probability that DENSITY = c divided by the probability that DENSITY = 0. For example, the coefficient −1.80 for WEIGHT in Table 2 implies that a unit increase in weight category is associated with a decrease in the ratio of the probability that DENSITY = 4 to the probability that DENSITY = 0 by a factor of exp(−1.80) = 0.165. The differing intercepts and coefficients for the various subcohorts imply, for example, that the ratio of the probability that DENSITY = 4 to the probability that DENSITY = 0 is greater in the subcohort of women given a biopsy examination in the screening phase and in the subcohort of women recommended for biopsy examination but not given a biopsy examination than in the normal subcohort.

For each subcohort, we estimated $p(x_1|x_2, I_t)$ from

$$p(x_1|x_2, I_t) = \sum_{i=1}^{8} p(x_1|x_2, t = i, I_t) p(t = i|x_2, I_t). \qquad [5]$$

To account for the fact that the distribution of DENSITY may differ in the three follow-up subcohorts (Table 1) and that "normal" women were underrepresented in the follow-up study,

Table 2. Estimated coefficients in the polytomous regressions for the conditional distributions of DENSITY given age and other covariates for the three follow-up subcohorts*

Cohort	DENSITY level	Covariate in polytomous regression models					
		Intercept	AGE	NUMREL	NBIOPS	AGEFLB	WEIGHT
Benign	1	1.08	−0.12	−0.01	0.77	0.52	−0.26
	2	2.32	−0.33	0.17	1.15	0.63	−0.75
	3	2.69	−0.63	0.36	1.66	0.93	−1.21
	4	2.10	−0.85	−0.05	1.88	1.07	−1.80
Normal	1	2.23	−0.08	0.26	0.00	0.20	−0.47
	2	3.00	−0.19	0.11	0.36	0.38	−0.93
	3	3.92	−0.49	0.37	0.64	0.60	−1.47
	4	2.42	−0.60	−0.06	0.88	0.95	−1.87
Recom	1	3.26	−0.09	−0.81	0.30	0.04	−0.49
	2	4.68	−0.32	−0.95	0.82	0.37	−1.06
	3	5.69	−0.56	−0.57	0.95	0.53	−1.56
	4	4.62	−0.65	−1.26	1.03	0.48	−1.75

*Covariates are coded as follows: age (AGE) was coded as 0, 1, 2, 3, 4, 5, 6, or 7 for a woman in age intervals <45, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, or ≥75 years at entry into the follow-up phase in 1980. The number of affected mother or sisters (NUMREL), number of previous benign breast biopsy examinations (NBIOPS), age at birth of first live child (AGEFLB), and weight (WEIGHT) were coded as in Table 1. Recom refers to the subcohort of women who were recommended for biopsy examination during the screening phase but did not receive it. DENSITY refers to mammographic density.

we combined the subcohort-specific estimates by reweighting them to obtain

$$p(x_1, x_2|I_t) = p(x_2|I_t) \times \sum_{\text{subcohorts}} \left[\sum_t p_{\text{subcohort}}(x_1|x_2, t, I_t) p(t|x_2, I_t) \right] \times p(\text{subcohort}|I_t), \quad [6]$$

where $p(\text{subcohort}|I_t)$ refers to proportions of white women in age stratum I_t in each of the three follow-up subcohorts in the entire BCDDP population at the end of the screening phase. These proportions were 0.099, 0.868, and 0.032, respectively, for the benign breast disease, normal, and recommended-for-biopsy subcohorts among those younger than 50 years; corresponding proportions for those 50 years old or older were 0.082, 0.885, and 0.033. The joint mass distribution $p(x_1,x_2|I_t)$ from Eq. 6 was substituted into Eq. 3 to estimate $\text{AR}(I_t)$, which was substituted into Eq. 2 to obtain $h_1(t)$.

Estimating Absolute Risk and Its Variance

We calculated absolute risk of invasive breast cancer from age a to age $a + \tau$, where τ is the duration of the risk projection interval, in a woman with relative risk $r(x)$ from formula 5 of Gail et al. *(1)*:

$$\varphi(x) = \int_a^{a+\tau} [1 - \text{AR}(u)] r(x) h_1^*(u) \times \exp\left[- \int_a^u \{[1 - \text{AR}(v)] r(x) h_1^*(v) + h_2(v)\} \text{d}v \right] \text{d}u. \quad [7]$$

To compute the absolute risk from Eq. 7 and its variance, we assumed that the hazards $h_1^*(t)$ and $h_2(t)$ were constant on 5-year intervals as in Eq. 6 in Gail et al. *(1)*. To compute the variance of the estimate of $\varphi(x)$, we took into account random variation and covariation of estimates of the following quantities: the log relative risk parameters β, the log odds ratio parameters $\gamma = (\gamma_{0c}, \gamma_c, \delta_c)$, where $c = 1, 2, 3$, or 4, to model $p_{\text{subcohort}}(x_1|x_2,t)$ in each subcohort, $p(t|x_2,I_t)$, and $p(x_2|I_t)$. Because values of β and γ were estimated by use of data from the BCDDP but values of $p(t|x_2,I_t)$ and $p(x_2|I_t)$ were from the NHIS, their estimates were independent. However, estimates of β and γ were correlated for two reasons. First, most of the control subjects from the screening phase were included in the normal subcohort, and a large fraction of them remained healthy and were, therefore, used to estimate γ. Consequently, there was correlation between the contribution to the estimate of β from the screening phase and γ estimated from the women in the normal subcohort who remained free of disease. Second, there was a correlation between the contribution to the estimate of β from each subcohort and the estimate of γ from the same cohort. Estimates of $p(t|x_2,I_t)$ and $p(x_2|I_t)$ were correlated as well. Confidence limits on the estimated absolute risk were calculated from the variance of the estimate of $\varphi(x)$ by using the delta method to compute the variance of $\log[\varphi(x)/(1 - \varphi(x)]$, computing symmetric two-sided 95% confidence intervals on the logit-transformed risk, and back-transforming the result. The detailed variance calculations are presented in the Appendix. Numerical analysis that was based on formulas in the Appendix revealed that the largest contribution to the variance of $(1 - \text{AR})$, and hence to the variance of absolute risk for a woman at the baseline level of risk factors, is from the variance of $\hat{\beta}$. Indeed this source contributes at least 20 times more than any other source to the variance of $(1 - \text{AR})$.

To provide the reader with an approximate confidence interval on $\varphi(x)$, we regressed the upper confidence limits on $\varphi(x)$, which were calculated from the variance estimates in the Appendix, on $\varphi(x)$ and $\varphi^2(x)$ (upper locus in Fig. 2). A similar regression analysis was obtained for the lower confidence limits (lower locus in Fig. 2). The points used in the regressions were chosen to cover a broad range of absolute risks as follows. For each of the 14 starting ages of 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, and 85 years, we considered projection intervals with a length of 5, 10, 20, or 30 years, which were subject to the constraint that the starting age plus the duration of the projection interval was no greater than 90 years. This yielded 47 possible age intervals over which projections were to be made. For each such age interval, we computed the absolute risk for each of the 1080 possible risk factor combinations, resulting in $47 \times 1080 = 50\,760$ pairs (upper confidence limit, absolute risk) and 50 760 pairs (lower confidence

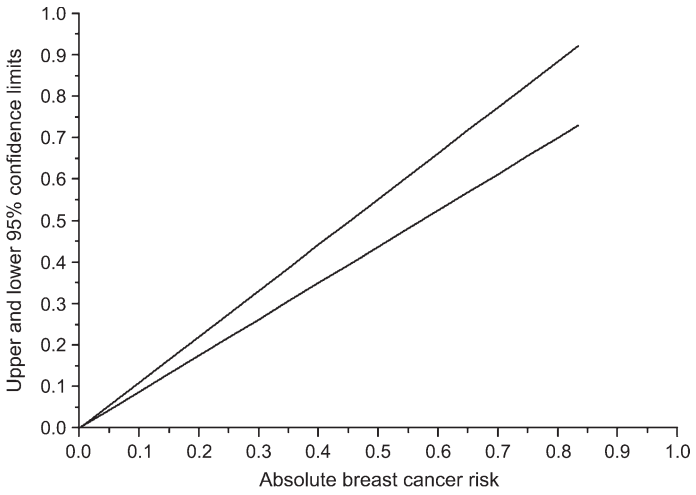

Fig. 2. Upper and lower 95% confidence limits for absolute risk from the new model with mammographic density.

limit, absolute risk). Thus, each of these quadratic regressions was based on 50 760 points. The squared multiple correlations for the upper and lower confidence intervals were 0.9991 and 0.9989, respectively. Thus, the loci in Fig. 2 provide good fits to the calculated confidence limits and explain 99.9% of the variation in the confidence limits in these 50 760 scenarios.

Results

Relative Risk

We tabulated the relative risk factors from Eq. 1 (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results) to facilitate computation of the combined relative risk. As described previously *(1,10)*, we included an adjustment factor F for information on atypical hyperplasia (Table 1). As an example, we consider a woman with AGEFLB = 2, NBIOPS = 1, NUMREL = 1, WEIGHT = 3, and DENSITY = 2, and we assume it is unknown whether atypical hyperplasia was present on the biopsy (factor F). From Table 1, her combined relative risk is calculated as $1.37 \times 1.29 \times 1.56 \times 1.91 \times 1.95 \times 1.00 = 10.3$. To estimate the absolute risk for such a woman, we need to combine this relative risk with information on the age interval over which risk is to be projected (see Table 5), as described below.

Distribution of Covariates in the General Population, Attributable Risks, and Hazard Rates

The conditional distribution of DENSITY for age interval i, NUMREL, NBIOPS, AGEFLB, and WEIGHT is given by Eq. 4 with estimated parameters in Table 2 for each of the three BCDDP subcohorts. Density decreases with increasing WEIGHT and age interval i, as indicated by the negative coefficients (Table 2). As an example, consider a woman in the normal subcohort with AGEFLB = 1, NBIOPS = 0, and NUMREL = 1. If she is aged 51 years (age group $i = 3$) and weighs 160 pounds (WEIGHT = 3), then from Eq. 4 and data in Table 2, we calculate her chance of having DENSITY = 4 (i.e., ≥75% dense tissue) as $\exp[2.42 - (0.6 \times 3) - (0.06 \times 1) + (0.88 \times 0) + (0.95 \times 1) - (1.87 \times 3)]/(1 + 2.829 + 1.139 + 0.372 + 0.017) = 0.0031$. If instead this woman were in age group $i = 1$ (i.e., age 45–49 years) and had WEIGHT = 0, the probability of having DENSITY = 4 would be 0.1083, nearly 35 times as large.

Combining data on NUMREL, NBIOPS, AGEFLB, WEIGHT, and age with the conditional distributions in Table 2 and averaging as in Eq. 6, we estimated the joint distribution of all risk factors. The marginal distributions from this joint distribution are displayed for US women in Table 3. For comparison, we gave the marginal distribution among white control women with DENSITY data in the BCDDP screening phase. Examining the results for

Table 3. Estimates of the marginal distributions of covariates in percent in the general US population and in Breast Cancer Detection Demonstration Project (BCDDP) screening-phase control subjects

			Covariate values*					
Age, y	Target population	Risk factor code	0	1	2	3	4	5
<50	US women	DENSITY	10.5	28.1	25.6	29.5	6.2	
		NUMREL	94.5	5.3	0.2			
		NBIOPS	96.3	2.8	0.9			
		AGEFLB	18.0	22.5	51.0	8.5		
		WEIGHT	0.8	23.3	39.0	18.3	12.1	6.6
	BCDDP screening-phase control subjects	DENSITY	12.8	29.3	25.6	27.8	4.5	
		NUMREL	88.1	11.4	0.4			
		NBIOPS	87.0	10.2	2.6			
		AGEFLB	13.6	48.5	32.2	5.7		
		WEIGHT	1.1	30.1	44.2	16.0	6.3	2.2
≥50	US women	DENSITY	21.1	43.2	24.7	9.6	1.3	
		NUMREL	87.7	10.9	1.4			
		NBIOPS	85.2	10.5	4.3			
		AGEFLB	22.1	39.9	30.8	7.2		
		WEIGHT	0.9	16.0	36.9	25.4	14.3	6.5
	BCDDP screening-phase control subjects	DENSITY	14.2	39.5	29.6	14.2	2.4	
		NUMREL	86.0	12.8	1.2			
		NBIOPS	81.3	13.4	5.3			
		AGEFLB	9.0	32.6	45.2	13.2		
		WEIGHT	1.4	27.1	43.9	18.4	6.8	2.4

*The numerical codes are as defined in Table 1. Higher codes correspond to increased breast cancer risk compared with lower codes. Rows sum to 100%. DENSITY refers to mammographic density; AGEFLB refers to age at birth of first live child; NUMREL refers to number of affected mother or sisters; NBIOPS refers to number of previous benign breast biopsy examinations; WEIGHT refers to weight.

Table 4. Age-specific composite $h_1^*(t)$ and baseline $h_1(t)$ invasive breast cancer incidence rates and competing mortality rates $h_2(t)$ ($\times 10^{-5}$) for the Gail model 2 in the National Cancer Institute risk disk and for the new model

	Gail model 2*			New model†		
Age, y	$h_1^*(t)$	$h_1(t)$	$h_2(t)$	$h_1^*(t)$	$h_1(t)$	$h_2(t)$
20–24	1.0	0.6	49.3	1.1	0.3	42.3
25–29	7.6	4.4	53.1	7.6	1.7	47.0
30–34	26.6	15.4	62.5	25.2	5.6	62.9
35–39	66.1	38.3	82.5	62.5	13.8	90.7
40–44	126.5	73.2	130.7	124.3	27.5	131.3
45–49	186.6	108.0	218.1	204.7	45.2	192.4
50–54	221.1	128.0	365.5	278.1	70.2	311.2
55–59	272.1	157.5	585.2	353.0	89.2	522.4
60–64	334.8	193.8	943.9	408.1	103.1	873.8
65–69	392.3	227.1	1502.8	463.1	117.2	1389.4
70–74	417.8	241.8	2383.9	509.2	128.8	2213.4
75–79	443.9	256.9	3883.2	528.5	133.7	3565.6
80–84	442.1	255.9	6682.8	500.5	126.6	6081.2
≥85	410.9	237.8	14 490.8	420.1	106.3	14 426.0

*The Surveillance, Epidemiology, and End Results (SEER) rates $h_1^*(t)$ were for white women from January 1, 1983, through December 31, 1987. The competing mortality rates $h_2(t)$ were from the National Center for Health Statistics from January 1, 1985, through December 31, 1987.

†The SEER rates $h_1^*(t)$ were for white women from January 1, 1996, through December 31, 2000. The competing mortality rates $h_2(t)$ were from the National Center for Health Statistics from January 1, 1996, through December 31, 2000.

US women, we note that 9.8% of women younger than 50 years old have no dense tissue, compared with an estimated 21.7% of older women. As anticipated, the proportion of women with affected first-degree relatives and the proportion of women with biopsies increased with age. The BCDDP screening-phase control subjects tended to have higher levels of these risk factors than women in the general US population. In particular, BCDDP screening-phase control subjects had higher proportions with affected first-degree relatives, with biopsy examinations, and with delayed age at birth of first live child. BCDDP screening-phase control subjects 50 years old or older also had higher proportions with dense breast tissue.

From Eq. 3, we estimated the attributable risk (i.e., AR) from the joint distribution of risk factors and the relative risk model (Eq. 1). Attributable risks of breast cancer associated with breast density, AGEFLB, NUMREL, NBIOPS, and WEIGHT were 0.779 (95% CI = 0.733 to 0.819) and 0.747 (95% CI = 0.702 to 0.788), respectively, for women younger than 50 years and 50 years or older. The quantity (1 − AR), which is needed in Eq. 2, was estimated to be 0.221 (95% CI = 0.181 to 0.267) among women younger than 50 years old and 0.253 (95% CI = 0.212 to 0.298) among those 50 years old or older. The corresponding values of (1 − AR) in Gail model 2 *(2)* were 0.579 and 0.579. Thus, the attributable risk in the new model that included DENSITY is considerably larger than the attributable risk in the model with standard risk factors only. This difference accounts for the fact that the baseline hazard rates $h_1(t)$ are considerably lower in the new model (Table 4) than in the Gail model *(2)*. Table 4 also displays the hazards of competing causes of mortality $h_2(t)$.

Absolute Risk of Invasive Breast Cancer

From the piecewise constant hazards version of Eq. 7 and the hazards in Table 4, we calculated the absolute risks of invasive breast cancer for various initial ages at counseling, follow-up intervals, and initial relative risks (Table 5). First, we assumed that the relative risks determined at the time of initial counseling remained constant over the risk projection interval. For example, if we consider a 50-year-old woman with a relative risk of 3.0, from

Table 5. Projected absolute risks in percent for various initial relative risks (RRs), initial ages, and ages at the end of the projection interval

		Absolute risk in percent						
Starting age, y	Ending age, y	RR = 1.0	RR = 2.0	RR = 3.0	RR = 5.0	RR = 10.0	RR = 20.0	RR = 30.0
20	25	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	30	0.0	0.0	0.0	0.0	0.1	0.2	0.3
	40	0.1	0.2	0.3	0.5	1.1	2.1	3.1
	50	0.5	0.9	1.4	2.3	4.5	8.8	12.9
	90	3.6	7.1	10.4	16.6	30.1	50.3	63.9
30	35	0.0	0.1	0.1	0.1	0.3	0.6	0.8
	40	0.1	0.2	0.3	0.5	1.0	1.9	2.9
	50	0.5	0.9	1.4	2.2	4.4	8.7	12.7
	60	1.2	2.4	3.6	5.9	11.4	21.6	30.5
	90	3.6	7.1	10.4	16.6	30.2	50.4	64.1
40	45	0.1	0.3	0.4	0.7	1.4	2.7	4.0
	50	0.4	0.7	1.1	1.8	3.5	7.0	10.2
	60	1.1	2.2	3.3	5.5	10.7	20.2	28.7
	70	2.1	4.1	6.1	10.0	19.0	34.3	46.5
	90	3.5	6.9	10.2	16.4	29.8	49.9	63.5
50	55	0.3	0.7	1.0	1.7	3.4	6.7	9.9
	60	0.8	1.6	2.3	3.8	7.5	14.5	20.9
	70	1.8	3.5	5.2	8.5	16.3	29.8	41.1
	80	2.7	5.4	8.0	12.9	24.0	41.8	55.1
	90	3.2	6.4	9.4	15.1	27.6	46.9	60.4
60	65	0.5	1.0	1.5	2.5	4.9	9.6	14.0
	70	1.0	2.1	3.1	5.1	9.9	18.8	26.8
	80	2.1	4.1	6.0	9.8	18.6	33.5	45.4
	90	2.6	5.1	7.5	12.2	22.7	39.6	52.3
70	75	0.6	1.2	1.8	3.0	5.9	11.4	16.6
	80	1.1	2.3	3.4	5.6	10.9	20.5	28.9
	90	1.8	3.5	5.2	8.4	16.0	29.1	39.8

Table 5, her absolute risks to ages 55, 80, and 90 years are, respectively, 0.01, 0.08, and 0.094 (or 1%, 8%, and 9.4%).

As a further illustration, consider a 50-year-old white woman with AGEFLB = 0, NBIOPS = 1, NUMREL = 1, WEIGHT = 2, and DENSITY = 2. If we assume that it is not known whether atypical hyperplasia was found in her breast biopsy examination, then from Table 1, we estimate her relative risk as $1.00 \times 1.29 \times 1.56 \times 1.54 \times 1.95 \times 1.00 = 6.04$. To estimate her absolute risk over 10 years from Table 5, we interpolated between the entries for relative risks of 5 and 10. The resulting estimate is $3.8 + [1/(10 - 5)](7.5 - 3.8)(6.04 - 5) = 3.8\ 0.77 = 4.57\%$.

The 95% confidence interval for such a projection can be approximated from Fig. 2. The actual loci are $-0.0006 + 1.1649\varphi(x) - 0.0701\varphi^2(x)$ for the upper locus and $0.0004 + 0.8622\varphi(x) + 0.0157\varphi^2(x)$ for the lower locus. By substituting $\varphi(x) = 0.0457$, we obtain a 95% confidence interval of 0.040 to 0.052 for the absolute risk of 0.0457 in the previous example, in agreement with Fig. 2. The widths of the confidence intervals increase with increasing absolute risk (Fig. 2).

A Gauss program *(14)* has been developed that computes the estimates of absolute risk and corresponding confidence intervals exactly for any combination of risk factors and age interval over which risk is to be projected. Using this program, we found for the previous example that the relative risk was 6.03 and the exact absolute risk estimate was 0.046 (estimated 95% CI = 0.041 to 0.052). These exact calculations agreed well with the approximate values obtained by interpolation in Table 5 and from Fig. 2.

Comparison Between the New Model With Mammographic Density and Gail Model 2

We compared 5-year absolute risk estimates from the new model with those from Gail model 2, which is used in the NCI's Breast Cancer Risk Assessment Tool (http://www.cancer.gov/bcrisktool/), for the white control women with complete data on NUMREL, NBIOPS, AGEFLB, WEIGHT, and DENSITY in the screening phase of BCDDP. Data from 270 women who were selected from the 1744 control subjects with complete data are shown in Fig. 3. There were 270 unique relative risks among these 1744 control subjects, and one woman was randomly selected from each of the sets of women having these 270 unique relative risks. The equiangular line, on which the ordinate equals the abscissa, is presented in Fig. 3 to divide projections from the new model that exceed the projections from Gail model 2 (which are above the line) from those projections from the new model that are less than the projections from Gail model 2 (which are below the line). The new model tends to have higher projections than Gail model 2 when the projections from the new model are more than 0.01 and lower projections than Gail model 2 when projections are less than 0.01. Among the 444 women with projected 5-year risk of less than 0.01 according to the new model, the projections from the new model were lower than those from Gail model 2 in 285 (64.2%). Among the remaining 1300 women with projected risks of at least 0.01, the projections from the new model were lower than those from Gail model 2 in only 390 (30.0%). Similar patterns were observed in strata defined by age groups (unreported analyses).

To understand discrepancies between the new model and Gail model 2, one must consider the two relative risk functions and

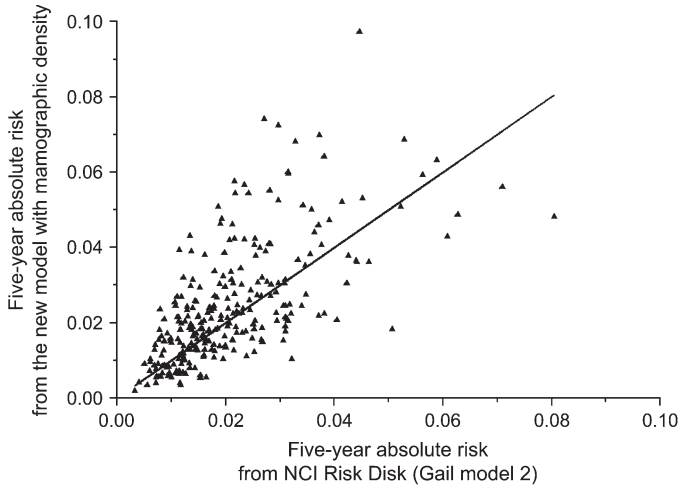

Fig. 3. Projected 5-year absolute risk of breast cancer from the new model with mammographic density (ordinate) plotted against the risk calculated with the National Cancer Institute's Breast Cancer Risk Tool (http://www.cancer.gov/bcrisktool/) (Gail Model 2). The data are from 270 screening-phase control subjects. There were 270 unique relative risks among the 1744 screening-phase control women, and one woman was randomly selected from each of the sets of women having these 270 unique relative risks. The **solid line** is an equiangular line, for which the value plotted on the ordinate equals the value plotted on the abscissa.

the fact that the baseline hazard $h_1(t)$ is lower in the new model than in Gail model 2 (Table 4). An inspection of the discrepancies between the new model and Gail model 2 indicated that the former tended to have higher risk projections for women in the higher categories of DENSITY and WEIGHT, as might be expected, because these factors are not present in Gail model 2. For example, the 5-year estimated risks for a 68-year-old woman with NBIOPS = 0, AGEMEN = 1, AGEFLB = 2, NULREL = 0, WEIGHT = 2, and DENSITY = 4 were 2.3% and 4.6%, respectively, for Gail model 2 and the new model; the corresponding combined relative risks were 1.70 and 7.98. The relative risk parameters for the standard risk factors in Gail model 2 differ from those in the new model, but direct comparisons are complicated because Gail model 2 includes interactions between NBIOBS and an indicator for an age of at least 50 years and between NUMREL and AGEFLB. Apart from the previous comments on DENSITY and WEIGHT and on high versus low estimated risks, we could not identify general classes of women for whom the new model gave predictably different absolute risk estimates from Gail model 2.

Discussion

To estimate absolute risk, we combined our previously developed relative risk model (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results), which includes mammographic density, with information on the distribution of risk factors from the NHIS and BCDDP studies, with SEER invasive breast cancer rates, and with national mortality hazard rates. We showed how to obtain an estimate of relative risk (Table 1) and use this relative risk to find the absolute risk estimate (Table 5). Variance calculations assume that SEER composite invasive breast cancer incidence rates and national mortality rates are known without error. However, these

calculations allow for random variation in the relative risk estimation, which is complicated by the patterns of missing data (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results); for the complex sampling in NHIS to estimate the joint distribution of age, AGEFLB, NBIOPS, NUMREL, and WEIGHT; and for estimation of the conditional distribution of DENSITY in the BCDDP subcohorts, given other risk factors (Table 2). These complex variance calculations are described in the Appendix. Approximate confidence intervals on absolute risk projections can be obtained from Fig. 2, and a Gauss program is available to compute the confidence intervals analytically.

The relative risk model used in this study was obtained by combining data from the BCDDP screening phase with data from the three follow-up subcohorts, even though there was evidence of heterogeneity of relative risk estimates between the screening phase and the follow-up phase (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). We used this strategy in the expectation that the resulting absolute risk projections would work reasonably well over both short-term and longer term intervals. Independent validation studies are needed to test this strategy. The calculations in the paper also assume that a woman's relative risk remains constant over the risk projection interval. If one anticipates that a factor such as NUMREL, DENSITY, or WEIGHT will change over time, one could adapt Eq. 7 by allowing the relative risk to vary with risk factors $X(t)$ that change with age t. For example, if risk factors remain constant, a 40-year-old woman with NBIOPS = 1, AGEFLB = 2, NULREL = 0, WEIGHT = 1, and DENSITY = 4 will have an absolute risk of 16.1% to age 70 years. If, instead, her DENSITY decreases to 3 and her WEIGHT increases to 2 at age 55 years, then her projected risk between ages 40 and 70 years would be 15.0%.

We found large attributable risks of 0.779 (95% CI = 0.733 to 0.819) for women who were younger than 50 years and of 0.747 (95% CI = 0.702 to 0.788) for women who were 50 years old or older. These attributable risks are larger than the values near 0.48 found for the Gail model *(1)*. The larger attributable risks with the new model reflect the high prevalence of women with DENSITY of more than 0 (Table 3) and the large relative risks associated with DENSITY (Table 1), which are factors that have been noted previously *(6)*.

Independent validation studies are needed to check the relative risks as well as the calibration *(2,5)* and discriminatory power *(5)* of the absolute risk model. A model is well calibrated if the observed numbers of breast cancer diagnoses in various subgroups of the population agree well with expected numbers based on the model *(15)*. We have indirect evidence that the new model is well calibrated. Previous studies have shown that Gail model 2 is well calibrated *(2,5)*. Even though predictions from Gail model 2 and the new model can differ appreciably (Fig. 3), the average predicted 5-year risk from the two models is quite similar for women in the following age intervals: younger than 45, 46–55, 56–65, and older than 65 years. When we analyzed data from the 1744 white control subject women in the screening phase with complete data on the standard risk factors DENSITY and WEIGHT, we found that the average predicted 5-year risks in these intervals were 0.8%, 1.5%, 2.2%, and 1.9%, respectively, for the new model and 0.7%, 1.5%, 2.1%, and 1.9% for Gail model 2. These data suggest that the new model gives average predictions close to those of Gail model 2 and is, therefore, well calibrated. Independent cohort data are needed to examine calibration in subgroups defined by factors such as age and levels of DENSITY and to check the model's applicability to nonwhite women. Ursin et al. *(16)* found similar relative risks for breast cancer associated with mammographic density among white, Asian American, and African American women.

The most widely used statistic to assess "discriminatory power" is the concordance statistic *(5,15)* or area under the receiver-operating curve. Unlike tests of calibration, which require cohort data, the concordance statistic can be estimated from a sample of women with breast cancer and from a sample of control subjects. By analyzing data from the case patients and control subjects from the screening phase of the BCDDP, Chen et al. (unpublished results) showed that the new model had greater age-specific concordance than Gail model 2. For example, for women aged 60–64 years, the concordance of Gail model 2 was estimated as 0.602, compared with 0.664 for the new model. The difference 0.063 had standard error 0.026, indicating a statistically significant increase in concordance (two-sided $P = .015$). Statistically significant increases in concordance were found for the age intervals of 45–49, 50–54, 60–64, and 70 years old or older, but increases were present in all seven age groups and ranged from 0.009 to 0.094. The unweighted average age-specific concordance for Gail model 2 was 0.596, compared with 0.643 for the new model. Thus, adding mammographic density improves the discriminatory power modestly compared to Gail model 2. Whether this improvement is sufficient to justify the additional expense of obtaining data on mammographic density may depend on the particular application of the risk projection model *(15)*. In any case, it would be useful to have additional independent data to compare the concordance statistic of the new model with that of Gail model 2.

Tice et al. *(17)* studied whether mammographic density, measured by the categories of the Radiology Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology, could add to the discriminatory power of the Gail model. They found that BI-RADS was an independent predictor of risk but that it barely increased the concordance above levels provided by the Gail model alone. For example, for women with complete information on Gail model covariates, the concordance was 0.70 from the Gail model alone and 0.71 with the addition of BI-RADS. One possible explanation for the smaller increase in concordance than found by Chen et al. (unpublished results) is that BI-RADS may be less informative than DENSITY. The difference may also be related to the fact that Tice et al. did not stratify their concordance calculations on age. Analysis of data from women of all ages results in higher concordance for the Gail model; this result may make it harder to detect the incremental value of BI-RADS.

Several limitations of our data and approach should be noted. As mentioned previously, to obtain relative risk estimates that might apply over shorter or longer time intervals, we combined data from the screening phase with data from the follow-up phase, even though estimates of relative risks for DENSITY and NUMREL were larger in the screening phase. Because DENSITY data were not available in NHIS, we needed to assume that the conditional distribution of DENSITY given age, AGEFLB, NBIOPS, NUMREL, and WEIGHT was the same in NHIS as in BCDDP to estimate the joint distribution of all

these variables in the general U.S. population. Because medical practices and diagnostic procedures have changed over time, relative risks that are based on BCDDP data from 1973 to 1995 may differ somewhat from current relative risks. In particular, our data on DENSITY were based on mammograms taken from 1973 through 1975 and were later assessed by a single evaluator, who outlined the dense areas *(9)*. Measurements that are based on newer technologies or that are made by multiple evaluators might yield different relative risks for DENSITY, although correlations among readings of the same films made by the evaluator in this study at different times were similar to correlations among different evaluators reported in other studies *(9)* (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). For these reasons, it is important that independent validation data be obtained to assess the relative risk features and absolute risk projections from this model.

A Gauss program is available to researchers from the corresponding author for further testing of this model. If the new model is shown to be valid in independent evaluations, a more convenient program could be made available for counseling and other applications.

Appendix

This appendix describes the statistical methods needed to compute confidence intervals on our estimates of absolute risk. This material documents our methods, and the approach may be useful to others who are developing risk models based on various data sources. The practitioner interested in using Tables 1 and 5 to project risk and Fig. 2 to approximate the confidence intervals can skip this Appendix.

We obtained the asymptotic formulas for calculating the variance of the absolute risk estimates by applying the delta method *(18)*. For density levels $c = 1, 2, 3$, or 4 in Eq. 4, let $\gamma^{\text{normal}} = (\gamma_{0c}^{\text{normal}}, \gamma_c^{\text{normal}}, \delta_c^{\text{normal}})$, $\gamma^{\text{benign}} = (\gamma_{0c}^{\text{benign}}, \gamma_c^{\text{benign}}, \delta_c^{\text{benign}})$, $\gamma^{\text{recom}} = (\gamma_{0c}^{\text{recom}}, \gamma_c^{\text{recom}}, \delta_c^{\text{recom}})$, and $\gamma = (\gamma^{\text{normal}}, \gamma^{\text{benign}}, \gamma^{\text{recom}})$. Further, let $p_{12c}^{l} = p_\gamma(x_1 = c|x_2, t, t < 50)$, $p_{12}^{l} = p(t = i|x_2, t < 50)$, $p_2^{l} = p(x_2|t < 50)$, $p_{12c}^{h} = p_\gamma(x_1 = c|x_2, t, t \geq 50)$, $p_{12}^{h} = p(t = i|x_2, t \geq 50)$, $p^{l} = (p_{12}^{l}, p_2^{l})$, and $p^{h} = (p_{12}^{h}, p_2^{h})$. The projected absolute risk φ is a function of (p^l, p^h), β, and γ. That is, $\varphi = \varphi\{\text{AR}_{t<50}[p^l, \gamma, r(\beta)], \text{AR}_{t\geq 50}[p^l, \gamma, r(\beta)], r(\beta)\}$. Taylor series expansion of φ around the true values of these parameters yields

$$\begin{aligned}\hat{\varphi} - \varphi \approx{}& \frac{d\varphi}{d\text{AR}_{t<50}} \frac{d\text{AR}_{t<50}}{dp^l}(\hat{p}^l - p^l) \\ &+ \frac{d\varphi}{d\text{AR}_{t\geq 50}} \frac{d\text{AR}_{t\geq 50}}{dp^h}(\hat{p}^h - p^h) \\ &+ \sum_{\text{three cohorts}} \left[\frac{d\varphi}{d\text{AR}_{t<50}} \sum_{c,t<50} \left(\frac{d\text{AR}_{t<50}}{dp_{12c}^l} \frac{dp_{12c}^l}{d\gamma^{\text{cohort}}} \right) \right. \\ &\left. + \frac{d\varphi}{d\text{AR}_{t\geq 50}} \sum_{c,t\geq 50} \left(\frac{d\text{AR}_{t\geq 50}}{dp_{12c}^h} \frac{dp_{12c}^h}{d\gamma^{\text{cohort}}} \right) \right] (\hat{\gamma}^{\text{cohort}} - \gamma^{\text{cohort}}) \\ &+ \left[\frac{d\varphi}{d\text{AR}_{t<50}} \frac{d\text{AR}_{t<50}}{d\beta} + \frac{d\varphi}{d\text{AR}_{t\geq 50}} \frac{d\text{AR}_{t\geq 50}}{d\beta} + \frac{d\varphi}{d\beta} \right] (\hat{\beta} - \beta). \end{aligned} \quad [A1]$$

The variance of the absolute risk estimate can be obtained from Eq. A1 in terms of the variances and covariances of the parameter estimates, which are identified by hats over the parameters and by substituting estimates of terms such as $d\varphi/d\text{AR}_{t<50}$.

To use Eq. A1 in this way, we calculated the following quantities in Eq. A1.

$$\begin{aligned}\frac{d\varphi}{d\text{AR}_{t<50}} ={}& -\int_a^{50} \left[\int_a^u rh_1^*(v)dv(1 - \text{AR}_{u<50}) - 1 \right] rh_1^*(u) \\ &\times \exp\left[-\int_a^u \{[1 - \text{AR}(v)]rh_1^*(v) + h_2(v)\}dv \right] du \\ &+ \int_{50}^{a+\tau} [1 - \text{AR}_{u\geq 50}]rh_1^*(u) \int_a^{50} rh_1^*(v)dv \\ &\times \exp\left[-\int_a^u \{[1 - \text{AR}(v)]rh_1^*(v) + h_2(v)\}dv \right] du, \\ \frac{d\varphi}{d\text{AR}_{t\geq 50}} ={}& \int_{50}^{a+\tau} \left[(1 - \text{AR}_{u\geq 50}) \int_{50}^u rh_1^*(v)dv - 1 \right] rh_1^*(u) \\ &\times \exp\left\{ -\int_a^u \{[1 - \text{AR}(v)]rh_1^*(v) + h_2(v)\}dv \right\} du, \\ \frac{d\varphi}{d\beta} ={}& \int_a^{a+\tau} [1 - \text{AR}(u)]xrh_1^*(u) \\ &\times \exp\left\{ -\int_a^u \{[1 - \text{AR}(v)]rh_1^*(v) + h_2(v)\}dv \right\} du \\ &- \int_{\text{a}}^{a+\tau} [1 - \text{AR}(u)]rh_1^*(u) \int_a^u [1 - \text{AR}(u)]xrh_1^*(v)dv \\ &\times \exp\left\{ -\int_a^u \{[1 - \text{AR}(v)]rh_1^*(v) + h_2(v)\}dv \right\} du, \end{aligned}$$

$$\frac{d\text{AR}_t}{d\beta} = (1 - \text{AR}_t)^2 \sum_x x\, p(x_1, x_2|t) r(x_1, x_2),$$

$$\frac{d\text{AR}_{t<50}}{dp_{12i}^l} = (1 - \text{AR}_{t<50})^2 p_{i2}^l p_2^l r(x_1, x_2),$$

$$\frac{d\text{AR}_{t<50}}{dp_{i2}^l} = (1 - \text{AR}_{t<50})^2 \sum_{x_1} p_{12i}^l p_2^l r(x_1, x_2),$$

$$\frac{d\text{AR}_{t<50}}{dp_2^l} = (1 - \text{AR}_{t<50})^2 \sum_{x_1} \left[\sum_{t\in i} (p_{12i}^l p_{i2}^l) \right] r(x_1, x_2),$$

and

$$\frac{dp_{12c}^l}{d\gamma^{\text{cohort}}} = \binom{x_2}{t} [I(c > 0) - p_{12c}^l] p_{12c}^l, \quad c = 0, 1, 2, 3, 4.$$

The formulas for $d\text{AR}_{t\geq 50}/dp^h$ are exactly the same as for $d\text{AR}_{t<50}/dp^l$ except for suitable changes of superscripts from l to h that index probabilities conditional on $t < 50$ and $t \geq 50$, respectively.

We obtained the variance of $\hat{\varphi}$ by computing the variance of the right-hand side of Eq. A1. We noted that $\text{cov}(\hat{p}^l, \hat{\gamma}) = \text{cov}(\hat{p}^h, \hat{\gamma}) = \text{cov}(\hat{p}^l, \hat{p}^h) = \text{cov}(\hat{\beta}, \hat{p}^h) = \text{cov}(\hat{\beta}, \hat{p}^l) = 0$ because different data sources were used for each estimate. To obtain $\text{cov}(\hat{p}_{i2}^l, \hat{p}_2^l)$, we estimated the covariance matrix C for $\hat{p}(t, x_2|t < 50)$ and $\hat{p}_2^l$ with the software program SUDAAN *(13)*, which appropriately takes into account the cluster sampling scheme for the NHIS. Then we obtained the variance–covariance matrix between $\hat{p}_{i2}^l$ and $\hat{p}_2^l$ by applying the delta method. The variance–covariance matrix between $\hat{p}_{i2}^h$ and $\hat{p}_2^h$ was obtained in a similar fashion.

We calculated the covariance between $\hat{\gamma}$ and $\hat{\beta}$, which were both estimated from the BCDDP data. Recall that $\hat{\beta}$ was obtained as the linear combination of relative risk estimates from the screening phase and the three follow-up subcohorts, $\hat{\beta} = \hat{w}_s\hat{\beta}_s + \hat{w}_{\text{normal}}\hat{\beta}_{\text{normal}} + \hat{w}_{\text{benign}}\hat{\beta}_{\text{benign}} + \hat{w}_{\text{recom}}\hat{\beta}_{\text{recom}}$, in which w denotes matrix weights and the subscript s indexes the screening phase. We calculated $\text{cov}(\hat{\beta}_{\text{normal}}, \hat{\gamma}^{\text{normal}})$, $\text{cov}(\hat{\beta}_{\text{benign}}, \hat{\gamma}^{\text{benign}})$, and $\text{cov}(\hat{\beta}_{\text{recom}}, \hat{\gamma}^{\text{recom}})$ as described below. We included 3093 of the 3146 screening-phase control subjects in the

follow-up study, for 3017 (97.5%) in the normal subcohort, 48 (1.6%) in the benign subcohort, and 28 (0.9%) in the biopsy-recommended subcohort. Thus, we treated cov($\hat{\beta}_s$, $\hat{\gamma}^{\text{benign}}$) and cov($\hat{\beta}_s$, $\hat{\gamma}^{\text{recom}}$) as zero and calculated cov($\hat{\beta}_s$, $\hat{\gamma}^{\text{normal}}$) as described below.

We present the method for calculating cov($\hat{\beta}_{\text{normal}}$, $\hat{\gamma}^{\text{normal}}$) as an example. Let n_{normal} be the total number of subjects in the normal cohort, λ be the baseline breast cancer hazard parameters in the piecewise exponential model as described previously (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results), and φ be the parameters in the polytomous regression model for the distribution of DENSITY conditioned on age and weight used in the EM algorithm (Chen J, Ayyagari R, Chatterjee N, Pee DY, Schairer C, Byrne C, et al.: unpublished results). Let $U^j_\beta, U^j_\lambda, U^j_\varphi$ be the score functions from the observed data log likelihood for the normal subcohort for β, λ, and φ, respectively. The vector of scores is given by Eq. 7 in Chen et al. (unpublished results). We then wrote the influence function for ($\hat{\beta}_{\text{normal}}$, $\hat{\lambda}_{\text{normal}}$, $\hat{\varphi}_{\text{normal}}$) as

$$\begin{pmatrix} \hat{\beta}_{\text{normal}} - \beta \\ \hat{\lambda}_{\text{normal}} - \lambda \\ \hat{\phi}_{\text{normal}} - \phi \end{pmatrix} = \begin{pmatrix} \frac{1}{n_{\text{normal}}} \sum U^j_{\beta\beta} & \frac{1}{n_{\text{normal}}} \sum U^j_{\beta\lambda} & \frac{1}{n_{\text{normal}}} \sum U^j_{\beta\phi} \\ \frac{1}{n_{\text{normal}}} \sum U^j_{\beta\lambda} & \frac{1}{n_{\text{normal}}} \sum U^j_{\lambda\lambda} & \frac{1}{n_{\text{normal}}} \sum U^j_{\lambda\phi} \\ \frac{1}{n_{\text{normal}}} \sum U^j_{\beta\phi} & \frac{1}{n_{\text{normal}}} \sum U^j_{\lambda\phi} & \frac{1}{n_{\text{normal}}} \sum U^j_{\lambda\phi} \end{pmatrix}^{-1} \times \begin{pmatrix} \frac{1}{n_{\text{normal}}} \sum U^j_{\beta} \\ \frac{1}{n_{\text{normal}}} \sum U^j_{\lambda} \\ \frac{1}{n_{\text{normal}}} \sum U^j_{\phi} \end{pmatrix} \equiv \begin{pmatrix} \sum \tilde{U}^j_{\beta} \\ \sum \tilde{U}^j_{\lambda} \\ \sum \tilde{U}^j_{\phi} \end{pmatrix},$$

where $U^j_{\beta\beta} = dU^j_\beta/d\beta$, $U^j_{\beta\lambda} = dU^j_\beta/d\lambda$, and so on. In this expression, the summations are over all members of the normal subcohort. The quantity $\tilde{U}^j_\beta$ is the vector of influences of individual j on the components of $\hat{\beta}$, and $\tilde{U}^j_\lambda$ and $\tilde{U}^j_\varphi$ are corresponding influence vectors for $\hat{\lambda}$ and $\hat{\varphi}$. The covariance matrix of ($\hat{\beta}^T$, $\hat{\lambda}^T$, $\hat{\varphi}^T$)^T can be estimated from the influences. For example, Var($\hat{\beta}_{\text{normal}}$) can be estimated as $\sum(\tilde{U}^j_\beta)(\tilde{U}^j_\varphi)^T$ and Cov($\hat{\beta}_{\text{normal}}$, $\hat{\varphi}_{\text{normal}}$) can be estimated as $\sum(\tilde{U}^j_\beta)(\tilde{U}^j_\varphi)^T$, where the summations are over all n_{normal} members of the subcohort. If m_{normal} is the number of noncase patients in the normal subcohort who have DENSITY data, then the influence function for $\hat{\gamma}_{\text{normal}}$ can be obtained similarly:

$$\hat{\gamma}_{\text{normal}} - \gamma_{\text{normal}} = \left[\frac{1}{m_{\text{normal}}} \sum_{j=1}^{m_{\text{normal}}} U^j_{\gamma\gamma}\right]^{-1} \times \left[\frac{1}{m_{\text{normal}}} \sum_{j=1}^{m_{\text{normal}}} U^j_{\gamma}\right] \equiv \sum_{j=1}^{m_{\text{normal}}} \tilde{U}^j_{\gamma}.$$

As before, Var($\hat{\gamma}_{\text{normal}}$) can be estimated by $\sum(\tilde{U}^j_\gamma)(\tilde{U}^j_\gamma)^T$ and cov($\hat{\beta}_{\text{normal}}$, $\hat{\gamma}_{\text{normal}}$) can be estimated from $\sum(\tilde{U}^j_\beta)(\tilde{U}^j_\gamma)^T$, where now the summations are over the m_{normal} members of the normal subcohort who influence $\hat{\gamma}_{\text{normal}}$—namely, the noncase patients in the normal subcohort who have DENSITY data.

To calculate cov($\hat{\beta}_s$,$\hat{\gamma}_{\text{normal}}$), we noted that among the screening-phase control subjects who were included in the normal follow-up subcohort, 4.9% developed breast cancer and 95.1% remained healthy. Only the latter were used for the calculation of $\hat{\gamma}_{\text{normal}}$. Let n_{overlap} denote the number of individuals with nonzero influence on both $\hat{\beta}_s$ and $\hat{\gamma}_{\text{normal}}$. From the individual influences for $\hat{\beta}_s$—namely $\tilde{U}^j_{\beta s}$ obtained as described by Chen et al. (unpublished results), one can estimate cov($\hat{\beta}_s$,$\hat{\gamma}_{\text{normal}}$) as $\sum(\tilde{U}^j_{\beta s})(\tilde{U}^j_\gamma)^T$, where the sum is over the n_{overlap} individuals who influence both $\hat{\beta}_s$ and $\hat{\gamma}_{\text{normal}}$.

We derived the formula for the variance of the attributable risk similarly, as $AR_{I_t} = AR_{I_t}[p^l, \gamma^{\text{cohort}}, r(\beta)]$. The Taylor series expansion is

$$\hat{AR}(I_t) - AR(I_t) = \frac{dAR_{I_t}}{dp^l}(\hat{p}^l - p^l) + \frac{dAR_{I_t}}{d\beta}(\hat{\beta} - \beta) + \sum_{\text{cohort}} \frac{dAR_{I_t}}{d\gamma^{\text{cohort}}}(\hat{\gamma}^{\text{cohort}} - \gamma^{\text{cohort}}).$$

The variance can be obtained in a similar fashion as for the predicted risk. In fact, terms required to compute the absolute risk can be used for the calculation of the attributable risk.

References

(1) Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879–86.

(2) Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91:1541–8.

(3) Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. J Natl Cancer Inst 1998;90:1371–88.

(4) Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. J Natl Cancer Inst 1999;91:1829–46.

(5) Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001;93:358–66.

(6) Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995;87:1622–9.

(7) Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J Natl Cancer Inst 1995;87:670–5.

(8) Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. Radiology 2004;230:29–41.

(9) Benichou J, Byrne C, Capece LA, Carroll LE, Hurt-Mullen K, Pee DY, et al. Secular stability and reliability of measurements of the percentage of dense tissue on mammograms. Cancer Detect Prev 2003;27:266–74.

(10) Surveillance, Epidemiology, and End Results (SEER) Program SEER*Stat Database: Mortality—All COD, Total U.S. (1969–2000) with Expanded VA County Detail. National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released July 2003. Underlying mortality data provided by NCHS (http://www.cdc.gov/nchs). Available at http://www.seer.cancer.gov. [Last accessed: January 15, 2004.]

(11) Surveillance, Epidemiology, and End Results (SEER) Program. Public-use data (1996–2000), released 2003, based on November 2002 submission. Available at: http://www.seer.cancer.gov. [Last accessed: January 15, 2004.]

(12) National Center for Health Statistics. National Health Interview Survey (NHIS). Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. Available at: http://www.cdc.gov/NCHS/nhis.htm. [Last accessed: July 22, 2002].

(13) Research Triangle Institute. SUDAAN example manual, release 9.0. Research Triangle Park (NC): Research Triangle Institute; 2004.

(14) The Gauss System, version 6. Maple Valley (WA): Aptech Systems; 2005.

(15) Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. Biostatistics 2005;6:227–39.

(16) Ursin G, Ma H, Wu AH, Bernstein L, Salane M, Parisky YR, et al. Mammographic density and breast cancer in three ethnic groups. Cancer Epidemiol Biomarkers Prev 2003;12:332–8.

(17) Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res Treat 2005;94:115–22.

(18) Rao CR. Linear statistical inference and its applications. New York (NY): Wiley; 1965. p. 319–22.

NOTES

The National Cancer Institute has supported numerous studies that produced data on breast cancer risk factors and disease outcomes for women in the Breast Cancer Detection Demonstration Project, including, recently, data on mammographic density. The authors assume all responsibility for the proper use of these data, the analysis and interpretation of the data, the writing of the manuscript, and the decision to submit the manuscript for publication, subject to standard clearance procedures for manuscripts in the Division of Cancer Epidemiology and Genetics.

Manuscript received March 1, 2006; revised June 5, 2006; accepted July 20, 2006.